Editor in Chief Dr Wei WANG

International Journal of Security (IJS)

Book: 2009 Volume 3, Issue 4
Publishing Date:30-08-2009
Proceedings
ISSN (Online): 1985 - 2320

IJS Journal is a part of CSC Publishers
http://www.cscjournals.org

Published in Malaysia

Typesetting: Camera-ready by author, data conversation by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers

Table of Contents

Volume 3, Issue 4, July/August 2009.

Pages

Certification Authority Monitored Multilevel and Stateful Policy Based Authorization in Services Oriented Grids

Ajay Prasad ajayprasadv@gmail.com
School of Engineering
Sir Padampat Singhania University
Udaipur,313601, India

Saurabh Singh Verma ssverma.et@mitsuniversity.ac.in
Faculty of Engineering and Technology
Mody Institute of Technology and Science
Lakshmangarh, 332311, India

Ashok Kumar Sharma aksharma.et@mitsuniversity.ac.in
Faculty of Engineering and Technology
Mody Institute of Technology and Science
Lakshmangarh, 332311, India

Abstract

Services oriented grids will be more prominent among other kinds of grids in the present distributed environments. With the advent of online government services the governmental grids will come up in huge numbers. Apart from common security issues as in other grids, the authorization in service oriented grids faces certain shortcomings and needs to be looked upon differently. The CMMS model presented here overcomes all these shortcomings and adds to the simplicity of implementation because of its tight similarities with certain government services and their functioning. The model is used to prototype a State Police Information Grid (SPIG). Small technological restructuring is required in PKIX and X.509 certificates.

Keywords: Grid security, Authorization, Virtual Organization, Policy based authorization, Policy mappings.

1. INTRODUCTION

Grid Computing is a vastly distributed approach to facilitate solutions providing, resource sharing, job execution and various services. The grid term is derived from electricity grids. The analogy of electricity grid in compute grids is appropriately understood by the fact that both grids are mainly meant for supplying vastly distributed resources and services. The figures 1 and 2 depict the exact similarities between them. The Services oriented grids become more prominent as many government services are being diverted on the online platforms all over the world. The non-grid organization doesn't have connected nodes even within the organization. The user needing any service should enter the organization, go to the service site (or counter) and get the service. The organizational-bounded grids are confined to the organizational boundary. The user connects to the main node and gets its services by implicit forwarding to other internal nodes. It's understood that Virtualization is the main approach towards grids. A Virtual Organization (VO) based grids are those where services fall out of organizational boundaries, and multiple organizations jointly render services to users. The user connects to a discovery node nearby it and gets to a service node in order to get service. Figure 3 depicts organizations where some of their nodes are service nodes that take part in the grid. Grid Security has always been an issue. Providing services is not the only task that needs to be done. User's confidence and safety is a major concern and lots of energy is devoted into this factor too. Availability, integrity, confidentiality all are important and must be handled. Confidentiality consists of authentication and authorization. Authorization is to perform the necessary action to confirm that an authenticated entity is allowed to perform an action on a resource[14]. As stated in [13] there are two general approaches for authorization: identity-based or token-based. The identity based authorization mainly depends upon user identity and access control lists. On the other hand token based approaches rely on un-forgeable tokens granted by the service provider or controller which are presented by the users to avail services. Tokens of service can be passed on from one to another in form of trust management and delegation of rights. Trust[13] can be generally defined as having confidence that a party will behave in an expected manner despite the lack of ability to monitor or control that other party. In identity based authorization trust management consists of defining the sources of authorities for user identification, attribute assignment and possibly policy creation. Policies of an organization can be regarding access rights, levels of trust, whom to trust etc.

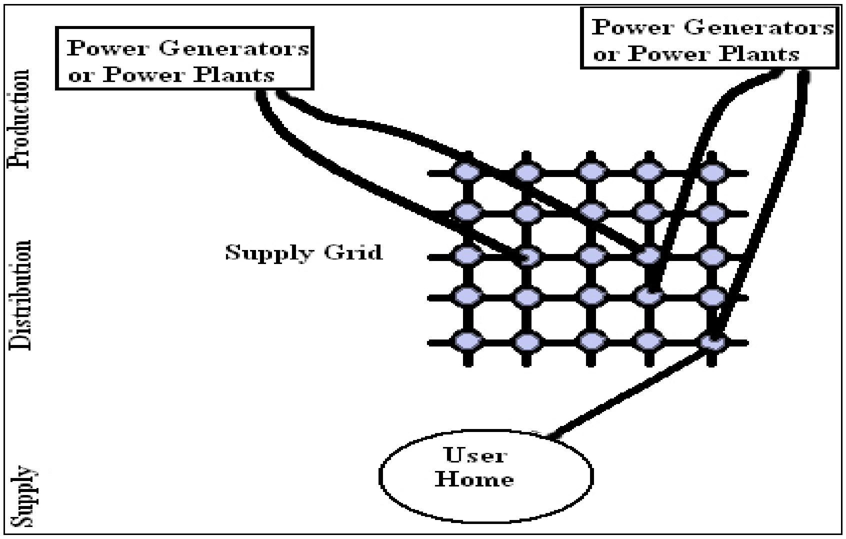

Power Generators or Power Plants
Power Generators or Power Plants
Production
Distribution
Supply
Supply Grid
User Home

Figure 1: An Electricity Power Grid.

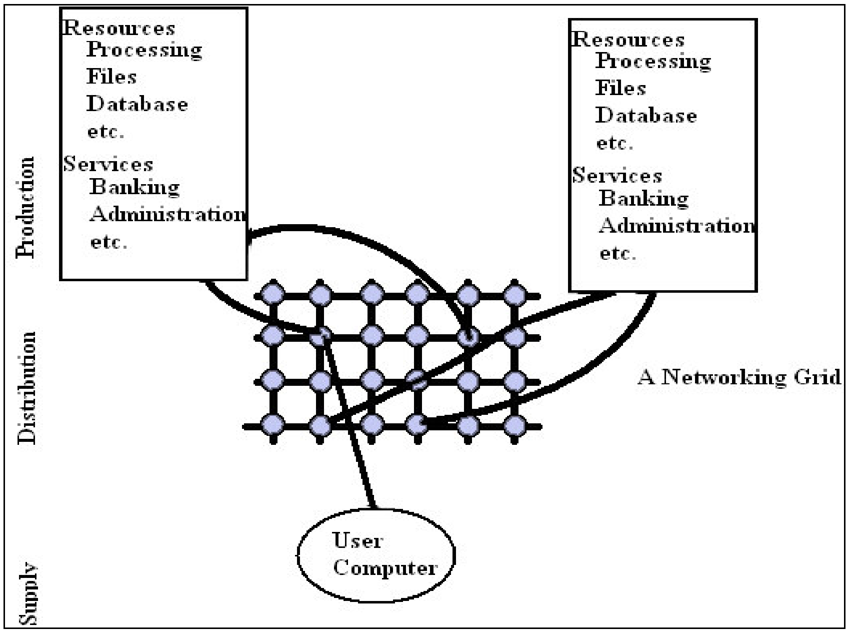


Figure 2: A Compute Grid.

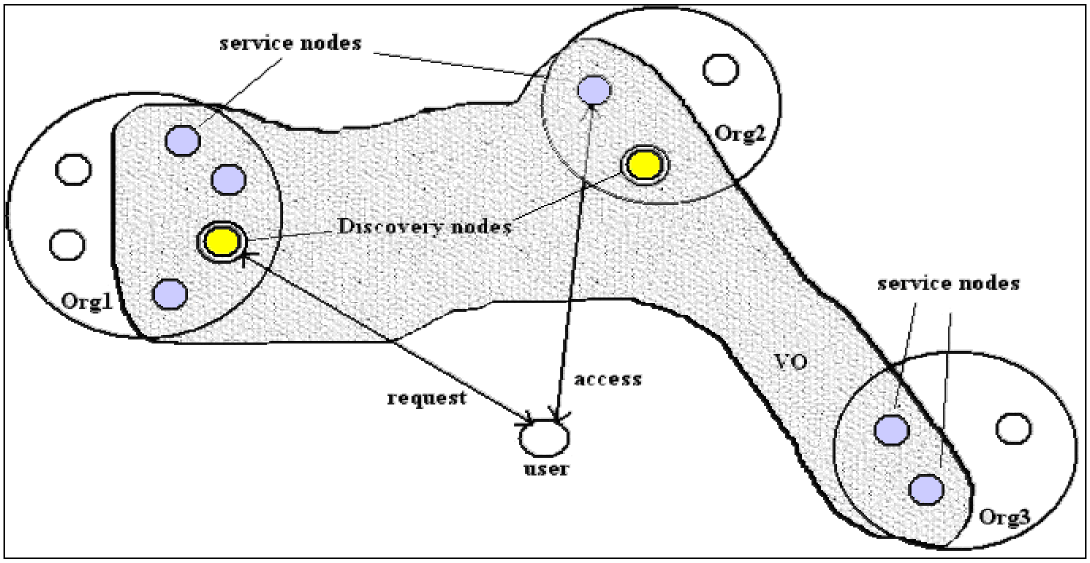


Figure 3: A Virtual Organization involving many organizations together.

2. AUTHORIZATION IN CURRENT GRID ENVIRONMENTS

Authorization has been taken care of by user identity and access control lists (distributed grid map files) in Grid Security Infrastructure (GSI)[1] incorporated in GT4.0. Also certain suggestions of a model where two points of action is required, namely, Policy Decision Point (PDP) and Policy Enforcement Point (PEP)[5]. The first one decides over the policy matters and the latter enforces the policies decided. The model can be realized both in simple and manageable centralized way or a complex but performance oriented distributed way. The governing rules regarding authorization either to grant or to deny can be acquired by the PDP in different ways[6] namely, Agent, Push and Pull models.

Trust model as suggested in [2] defines a policy database, evaluator and result of evaluation. It goes true only when dynamic, flexible and fine grained policies are maintained in a structured and simple manner. Even though dynamic policies are maintained as property based certificates in

PRIMA[7], VOMS[8], CAS[9] and X.509 attribute certificates[10] there is no standard interface for using them yet. Globus Resource Acquisition and Management (GRAM) system maintains policy mappings using gatekeeper and Job manager components[12]. Akenti Policy language[11] expresses policy in XML and store in three types of signed certificates: policy certificates, use-condition certificates and Akenti attribute certificates[11].

A proposed privacy, trust and policy based authorization framework[2] introduces the concept of Filter-in and Filter-out at domains in a distributed infrastructure. With Filter-out component, the Subject leaves the Domain with access rights that his parent Domain grants to him. With Filter-in component, the Subject enters the organization with access rights that the target Domain grants to the parent Domain. In other words, the Subject gets the intersection of the rights that his parent Domain grants to him and the rights that target Domain grants to parent Domain.

A Stateful grid should maintain state information along with certification to be carried during the lifetime of the process or job or request for resource. The safety and consistency in policy based authorisation systems are discussed in [4] suggest that the collection of credentials used to satisfy a given authorization policy acts as a partial snapshot of the system within which the policy is evaluated. The correctness of an authorization decision depends on the validity and stability of the view used during policy evaluation. If we assume that each credential is stable (i.e, that the assertion stated in the credential remains true until its pre-ordained expiration time) then policy evaluation can be reduced to the problem of stable predicate evaluation on distributed snapshots[3].

3. SHORTCOMINGS IN AUTHORIZATION TECHNIQUES IN CURRENT GRID ENVIRONMENTS

As mentioned in the above section, authorizations in grids are incorporated or suggested in many models or papers. They all have some of the following shortcomings in them.

Weak trust management: In diverse and dynamic domain to domain interactions trust management helps in deciding what entities are to be trusted to do what actions [13]. The trust management itself is a weakness in grid environments. However, large distributions of services have to rely on trust to some extent. Trust based on mere confidence can lead resources and services in wrong and malicious hands and can corrupt vital resources and their accessing.

Policies are static: Domain specific policies are essential so as to integrate multiple policies and varied policies of entities providing services and resources in a grid. Currently users must keep track of the right/required credential for each resource they might access.

Multi-nodal policies are not considered: Each entity or domain in a grid can be seen as a node, every node can have its own policies. Since a request can be serviced by combination of several nodes, their composite policies have to be considered or at least the requester must fulfil policy requirement of all the nodes once before getting started. This filtering is required from both sides (the requester and the service provider). This scenario can occur and can be fatal in case of job submissions in a grid (processing grids). If a job is submitted to a grid and one or more nodes later on reject one or more sub-jobs at the time of join then it could cause unnecessary delay and even losses.

State information of processes and/or users are not included in the authentication procedure or authentication exchanges: Keeping state information in authentication dialogues will help the policy management (supporting dynamic policies) to perform policy mappings[14] in more secure manner. In addition to the processes/requests the above goes true for any role based application working on a VO. The Dynamic Role Based Access control in SESAME[16] suggests a model but it is confined to role based access only.

Coarse grained policies and services: Coarse grained policies[14] results from the common reliance on standard operating system enforcement mechanisms. These are constrained by expressiveness limitations. It also poses an accounting problem if the user employs resources in conjunction with separate projects or applications in separate VOs. Coarse grained services result in scalability problems and need complex mechanisms for policy mappings.

4. THE CMMS MODEL

At the outset we need to take into consideration the scenario where the proposed model will work and then things can be placed in their respective slots. The service oriented grid is a grid where user can plug in (or connect) to get desired service (only if they are authorized to). One node can also act as a user while forwarding a request to the next node. The Grid Security Infrastructure's delegation of rights and negotiation of trust[15,17] can be performed in these cases. When a user connects to the VO discovery node, the discovery node performs initial checks and looks for a free or available service node and sends the address of the node to the user. Let us list out all the players in this grid scenario as shown in figure 4.

1. Users.
2. VO discovery nodes.
3. VO service nodes.
4. Certification authority.

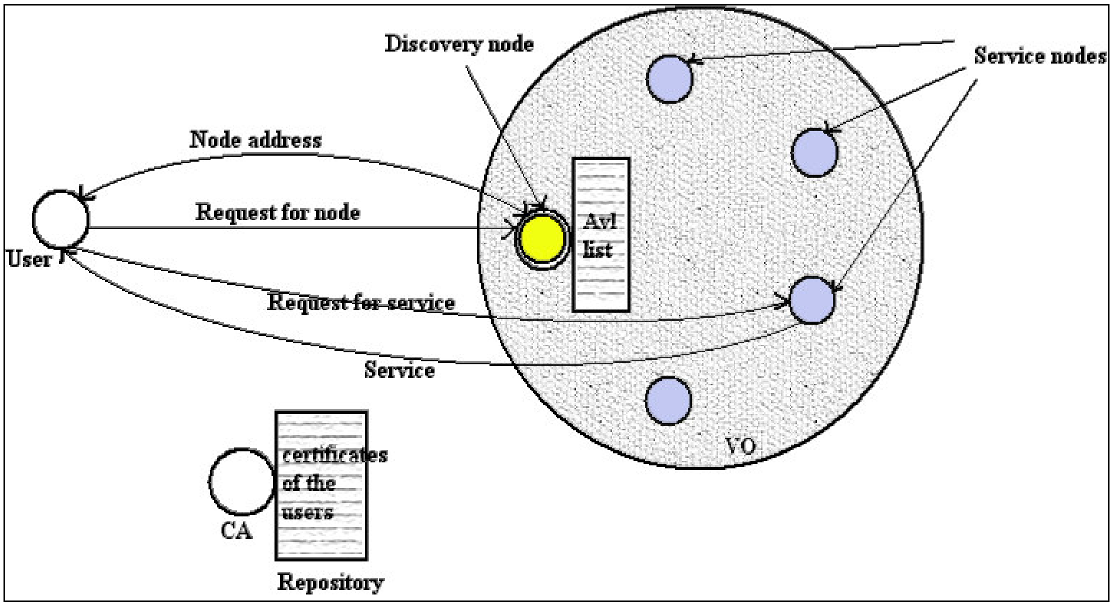


Figure 4: A normal service oriented grid scenario.

The users are the ones who may need to get the services of any node in the VO. The VO discovery nodes are responsible to perform initial authentication and give the access to an available node (address). The users then connect to the service node directly and get the required service. In all these dialogues the important aspect of authentication is taken care of by the certificates maintained by CA.

The proposed model will be consisting of four set of applications apart from general middleware at grid, CA and user site. Let's discuss these applications as we come across them one after another. The modified parts are put into the grid scenario in the figure 4 as figure 5. Major Players in the scenario are:

1. User application
2. Discovery node
3. Service node
4. CA node
5. Repository node
6. Monitor node

The user has an application in its system. The application will be basically of two parts or modules:

1. Get access part
2. Services interface part.

The get access part is responsible to perform connections and authentications and request for authentications both to VO discovery node and to the services nodes. This part is also responsible for requesting certificates of VO discovery nodes and service nodes to start authentication process. The get access part on receiving the service list activates the service interface part. Only those interfaces are initialised and activated which are in the service list that is sent by the service node. In case of forwarding of request (delegation of rights) the service list have to be matched at the next node and policy will be mapped once again and services need to be rendered if the mapping too has the user authorized for the service that has desired the request forwarding.

The discovery node is responsible to perform level1 authentication. It will perform Filter and Impose on the user states in the certificates. The discovery node maintains a list of available or free nodes. The discovery node selects one free node and sends it back to the user node.

The node address is sent to the user and the user get-access module connects to the service node. The service node contains two modules:

1. Authentication and Authorization module.
2. Service renderer module.

The Authentication and Authorization module is responsible to perform level 2 authentication demanding user's certificate from CA repository node and then perform policy mapping from the user's effective state sent by the discovery node. The policy mapping is explained in the later section.

The monitor node performs state or policy changes for users and service nodes. It sends signals to CA for policy changes and CA writes new certificate and sends it to the repository node which performs store cert to store it. The service renderer module will also perform service request forwarding and delegation of rights for the user.

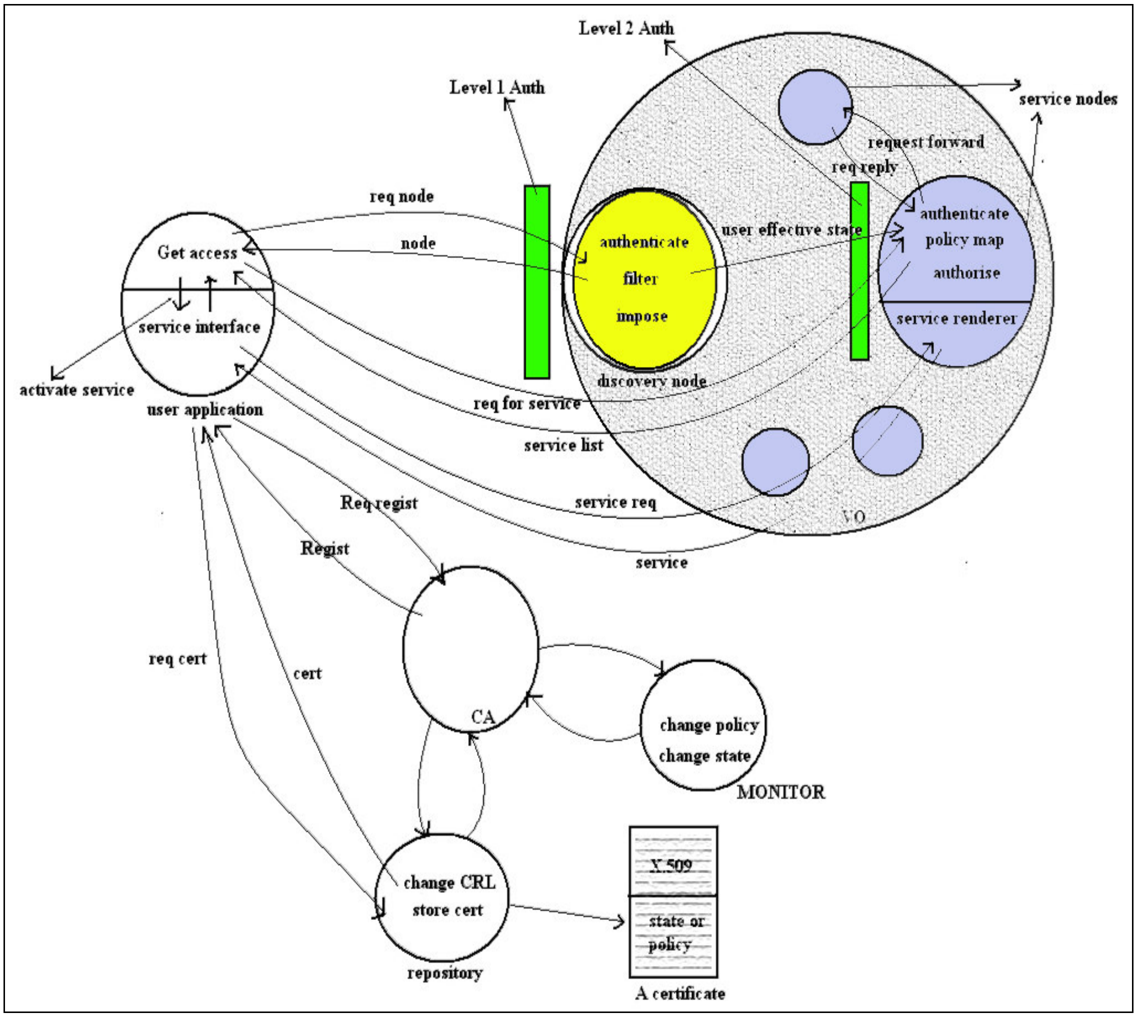


Figure 5: The CMMS overall scenario

Let us now have a look at the complete flow of the scenario as depicted in figure 5.

1. REG_USER: The user registers with the CA. The CA consults the monitor for initialising the state of the user.

2. REG_DISC, REG_SERV: The VO discovery node and service nodes are all certified similarly.

3. The user now installs the user module and invokes it and connects to the discovery node.

4. GET_CERT: The user requests CA repository for discovery node certificate, verifies it and extracts the node's public key and sends GET_NODE to discovery node.

5. The discovery node requests CA repository for the user's certificate, verifies it and extracts user's public key. It then performs AUTHENTICATE, FILTER and IMPOSE.

6. SEND_NODE: The discovery node selects a free service node and sends its address to the user.

7. SEND_EFF_STATE: The discovery node simultaneously sends the effective state of the user with the user name to the service node.

8. SERV_REQ: The user on receiving the service node address connects to the service node and sends service request message to the service node.

9. The service node on receiving the SERV_REQ from the user, requests for its certificate (GET_CERT) and verifies it, extracts user's public key.

10. Then the service node performs POLICY_MAP and sends the authorised SERVICE_LIST to the user.
11. The user avails the service.

(All the communications are through X.509 authentication procedures. The SPIG implementation has not implemented the authentication procedure as its just a prototype.)

The Middleware architecture:

The overall model can be viewed in an architectural fashion suited for implementations. The figure 6 depicts the modules that must be included as per incorporating the CMMS model. The figure is quite self explanatory. and has not been explained in this paper.

X.509 certificates and PKIX

The PKIX components are added with a Monitor node and every X.509 certificate will hold extra space for keeping the policies and states of users and service node. The space either can be adjusted from the v3 extensions in X.509 or can be added as required. The adding of an extra component in PKIX and an extra space in the X.509 format will never affect the authentication procedure of X.509 at all.

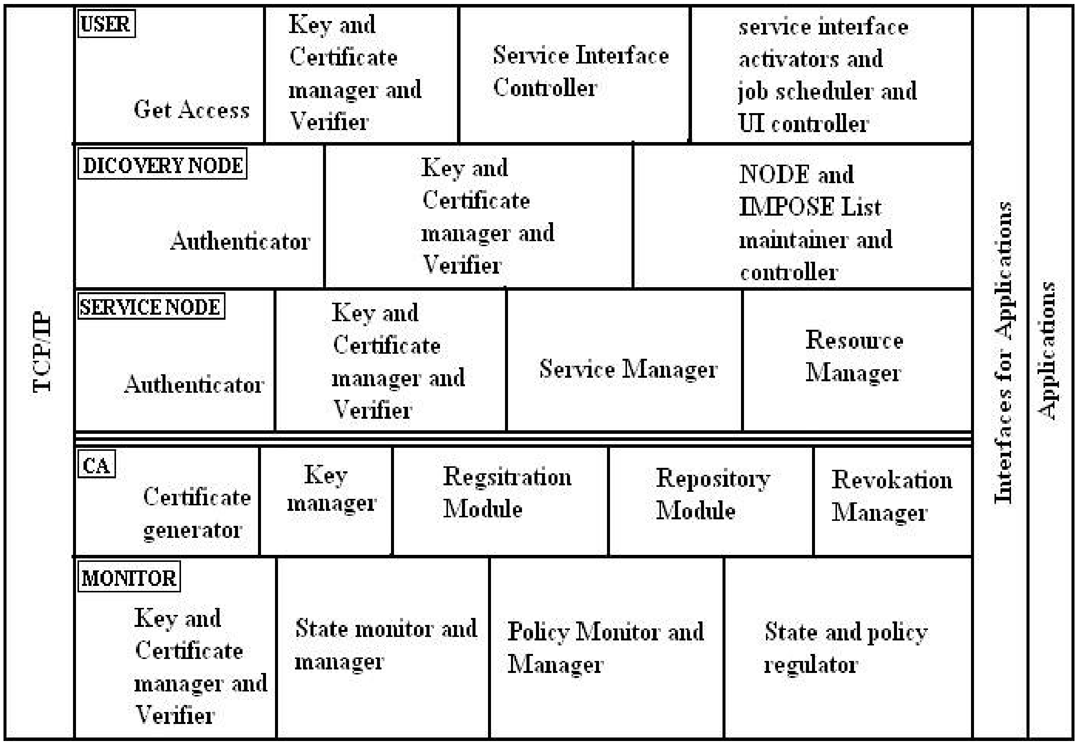

TCP/IP
USER
Get Access
Key and Certificate manager and Verifier
Service Interface Controller
service interface activators and job scheduler and UI controller
DICOVERY NODE
Authenticator
Key and Certificate manager and Verifier
NODE and IMPOSE List maintainer and controller
SERVICE NODE
Authenticator
Key and Certificate manager and Verifier
Service Manager
Resource Manager
CA
Certificate generator
Key manager
Regsitration Module
Repository Module
Revokation Manager
MONITOR
Key and Certificate manager and Verifier
State monitor and manager
Policy Monitor and Manager
State and policy regulator
Interfaces for Applications
Applications

Figure 6: The CMMS Middleware Architecture.

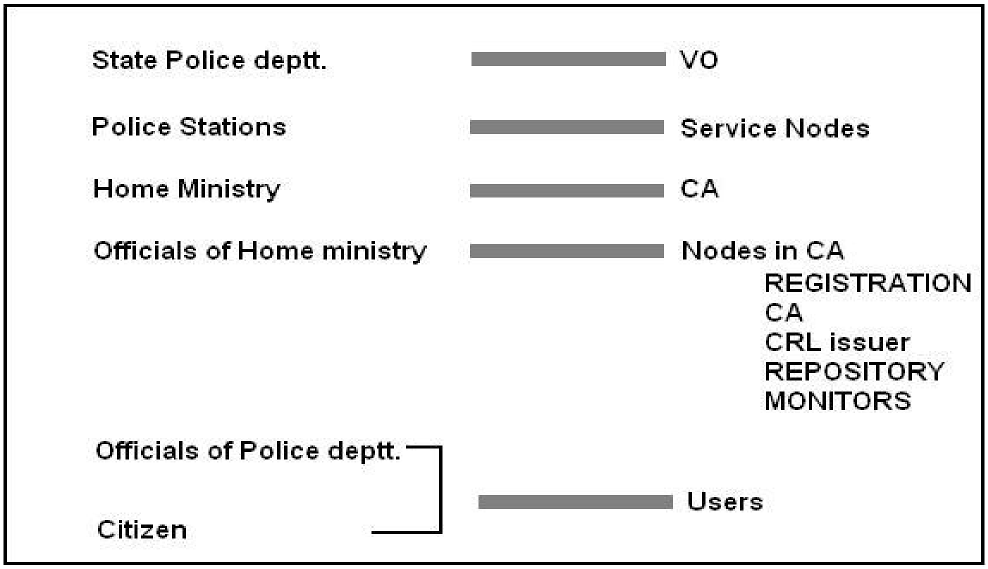
5. EXAMPLE: STATE POLICE INFORMATION GRID (SPIG) PROTOTYPE IMPLEMENTATION

The concept of CMMS is presented hereby through a prototype implementation called as State Police Information Grid (SPIG). The grid prototype is a service oriented VO keeping in mind the present Indian Police Department Infrastructure. The police in India are structured as police stations covering a specific area of control and authority. Every police station is responsible to take care or ensure law and order in their specified region. The activities of every police station are taken care of by number of officials ranging from Station in-charge to a set of inspectors. The groups of police stations are under the domain of Superintendents, Inspector generals etc. The overall policing of a state are governed by the state Home Ministry. The officials of Home Ministry range from political leaders to a set of IAS officers. The Home Ministry is responsible to monitor the police activities, regulate rules and policies, assign roles and perform authorized activities like transfers, suspension etc. The whole structure of policing as explained above might not be exact to the existing structure but that's not important. What's important for us is to understand the essence of CMMS with the structure.

In a nutshell the whole structure consists of following components.

1. State Police.
2. Police stations.
3. Home Ministry.
4. Officials of Home Ministry.
5. Officials of police.
6. Citizen of the state.

The whole CMMS can be realized with these components very simply. Lets map all the above components to the components of CMMS.



State Police deptt.	VO
Police Stations	Service Nodes
Home Ministry	CA
Officials of Home ministry	Nodes in CA REGISTRATION CA CRL issuer REPOSITORY MONITORS
Officials of Police deptt. Citizen	Users

Figure 7: State police structure mapped to the CMMS

The whole prototype is actually a set of clients and servers. The list below is the clients and servers implemented. They are not exactly the same number of servers as we'll have in CMMS. The later figure shows the mapping which reveals that actually the roles and tasks required in CMMS are all covered in the implemented set of servers.

The List of client and servers are:

1. USER
2. VO_SERV_NODE (SPIG)
3. VO_DISC_NODE (SPIG)
4. CA _NODE (CA)
5. CA _CRL_REP_NODE (CA)
6. CA _MONITOR_NODE (CA)

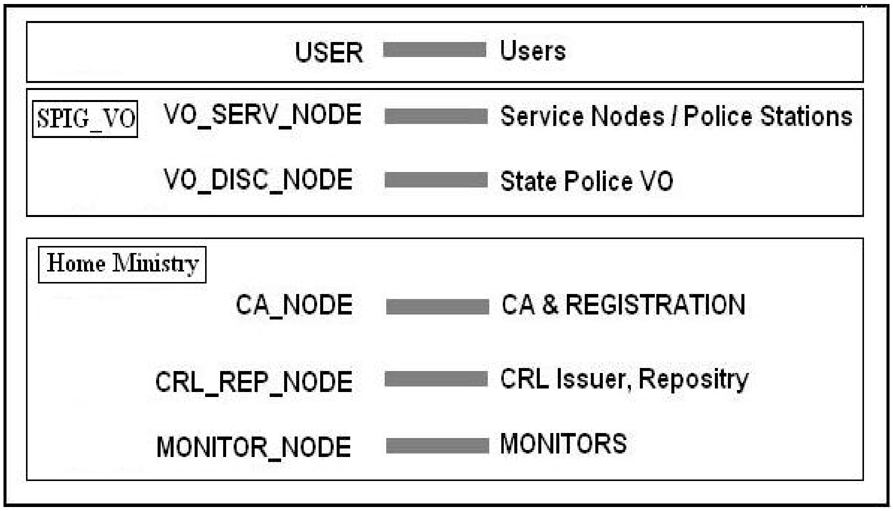


Figure 8: The implemented servers and their CMMS counterparts.

6. STATE TO POLICIES AND POLICIES TO SERVICES MAPPING IN CMMS

1) State in CMMS:

In most of the earlier authorization in grids more emphasis was given to roles and privileges. But, a stateful grid must take states of its user as well as servicing nodes into consideration while authorization. Here are few reasons why states are important in grid authorization.

1. Both users and the service nodes are subject to change (for users it can be personality or position and for service nodes it can be load and vulnerability).
2. Different users might have different policies as to how and when they'd like to take help from.
3. Different servicing nodes might have different policies regarding rendering of services to a user in a particular state.
4. It is important that the audit records of the services or resources are maintained as to which user and in what various states they have been used.
5. In case of processing-grids process state play roles while joining or forking a process over multiple nodes.
6. Suspicious nodes or users can be traced out and stabilized.

Example: states in the SPIG prototype.

State	Ref No
On Duty	1

Suspended	2
Transferred	3
Convicted	4
On Leave	5
View Restricted	6
Edit Restricted	7
User	8

Table 1: States of users used in SPIG.

A user state can change and in that case many accesses he/she was authorized may become vulnerable and might require restriction imposed on the accessibility of the user. The states of a user are a fact due to some environmental or situational changes but are very important in authorization. As shown in the table 1, The states 1 to 7 are for State Police employees or officials. The states 4, 7 are for common man or citizen of state.

In the same manner the states of the VO and its service nodes are also very important (The VO and Service node states are not included in SPIG prototype). For example if a VO is a disowned company then the users might not like to use it in full confidence. If a service node is in a state of partial failure or under control of certain enemy organization, it will not be encouraging for the users to use it for any purpose. However, considering the major issues the user state is vital and has been touched in SPIG implementation.

The state structure can be as:

<state-list> -> {<state>, <state>, …}
<state> -> 1/2/…/n *for n numbers of states maintained.*

One can be having more than one state.

2) Filter and Impose at VO level:

Filtering: Certain states of user which are not needed in policy mappings in a VO needs to be separated from the set of states in the user certificates.

Imposing: The discovery node will maintain a IMPOSED LIST and will add/impose certain other states to the user based on the list on per user basis.

Example: Filtered States

States	States considered at present	Filtered states
1, 4,15	1,2,3,4,5,6,7,8,9,10,11,12	1, 4

Table 2: Example: Filtered states.

State 15 is not under consideration in the VO so only state 1, 4 will be used in mapping.

Example: imposed States

In user certificate	Filtered state	Imposed state	Effective state

1, 4, 15	1, 4	11, 12	1, 4, 11, 12

Table 3: Example: Imposed States.

VO can impose certain states to certain users as per the IMPOSED_LIST maintained at VO level. The states imposed will vary from user to user and can be dynamically maintained at VO level in a IMPOSED_LIST. The imposed states can be determined by certain management policies which may or may not be programmed.

3) Services:

Services can be numbered in a very fine grained way so that they can be rendered that way. That is, a service like Police verification can be read as well as upgraded and applied for. These can be given unique numbers as:

1. Police verification: search and read
2. Police verification: search an upgrade
3. Police verification: apply

With fine grained services and policies and states are can have better control over authentication and less vulnerable to attacks.

Example: Services in SPIG.

Services	**ref no.**
Criminal Records Database	1
FIR Records	2
Search for INV status	3
ADD FIR/Criminal Records	4

Table 4: Services Provided in SPIG.

4) Policies.

Policies are like rules to be followed in order to allow or disallow a user from getting a set of services and also accepting or rejecting a service from a node by the user. Generally policies are based on states like for example a typical policy can be:

7: 3,4

That is, the users having state 7 will be given service number 3 and 4 only.
There can be several policies for various permutations of states. The policy structure can be as:

<policy> -> {<state> : <service-list>}
<service-list> -> {<service>, <service>, …}
<service> -> 1/2/3…/m
m= number of service rendered at that node

Example: A policy used in SPIG.

{1: 1, 2, 3, 4}
{2: 2, 3}
{3: 2, 3, 4}
{4: }
{5: 2, 3, 4}
{6: 2}
{7: 2, 3}
{8: 2}

Table 5: Example: Policy of a node

5) State-Policy to service mapping.

Based on the policies and user's state the service nodes can decide as to which services it will provide to the user. At the VO discovery node effective user state will be packaged and sent to the service node for mapping.

A typical procedure of mapping at service nodes is explained below. The procedure is used in the SPIG implementation. The service node will tally the effective state sent by the VO level to the state in the user certificate and perform mapping.

Lets see how the policies can be represented. The SPIG uses an exact representation as shown below:

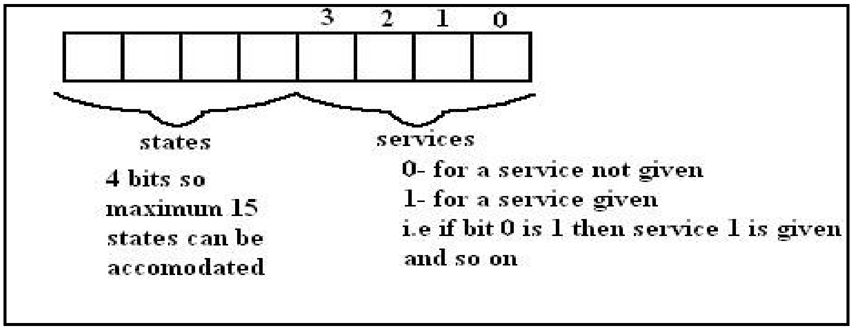
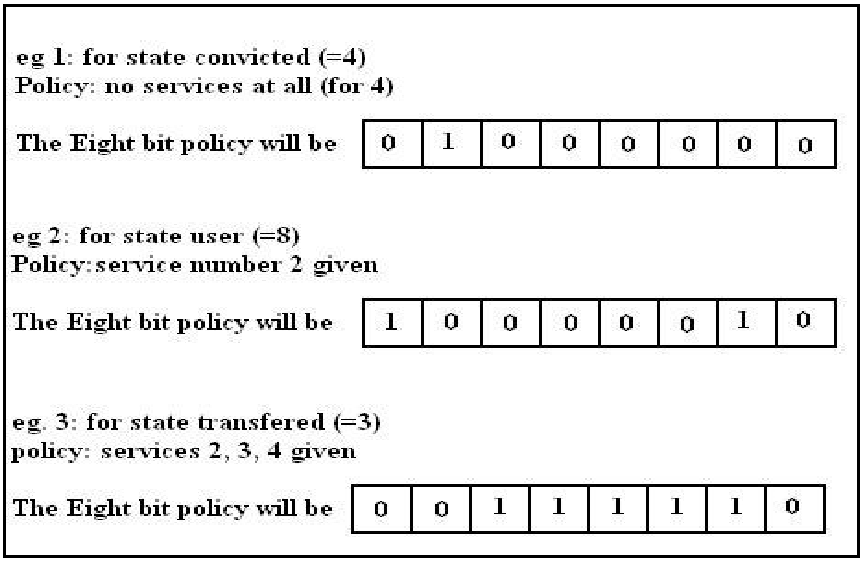


Figure 9a: Representing Policies



eg 1: for state convicted (=4)
Policy: no services at all (for 4)

The Eight bit policy will be

0	1	0	0	0	0	0	0

eg 2: for state user (=8)
Policy:service number 2 given

The Eight bit policy will be

1	0	0	0	0	0	1	0

eg. 3: for state transfered (=3)
policy: services 2, 3, 4 given

The Eight bit policy will be

0	0	1	1	1	1	1	0

Figure 9b: Example: keeping state and policy

Now, the mapping will be as below (using the policy stated earlier)

Say if a user has the state 5, 6
The mapping will be performed as follows

Goto 5th policy (94) = p1 = **01011110**
Goto 6th policy (98) = p2 = **01100010**

By masking get service for 5th policy
01011110
AND **00001111**

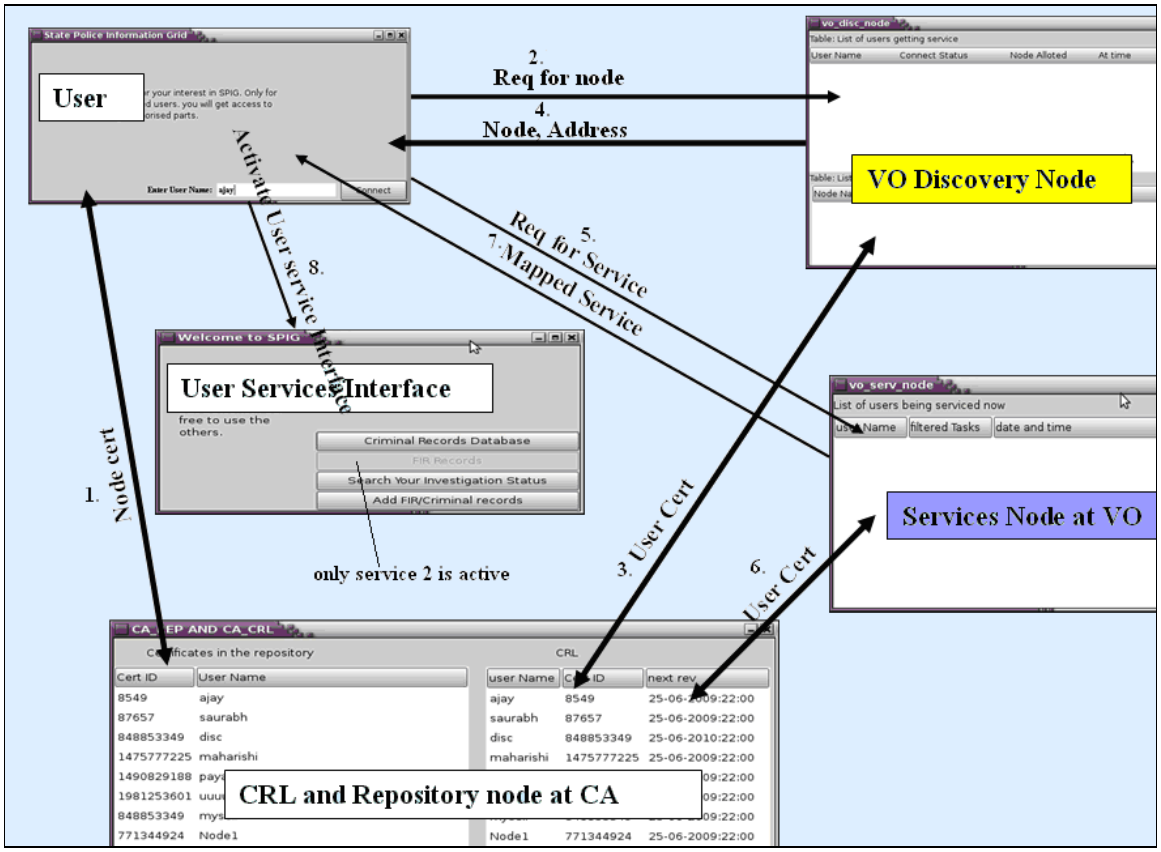
Service$_1$**00001110** =14
means given services 2,3 and 4
Similarily
Serivce$_2$**00000010** =2
means given service 2

Now, since there are more than one set of mapped services only those services will be rendered which are common in both the sets. Therefore we will AND both these services.

ANDing **00001110**
AND **00000010**

00000010 = 2
Therefore the mapped service is = 2
The user will thus get only 2nd service.

The same mapping technique is exclusively followed in SPIG implementation. Here is the working scenario in CMMS as implemented in SPIG. The numbers in the figure 10 depict the sequence in which the request or response will take place.



Cert ID	User Name
8549	ajay
87657	saurabh
848853349	disc
1475777225	maharishi
1490829188	paya
1981253601	uuu
848853349	mys
771344924	Node1

user Name	Cert ID	next rev
ajay	8549	25-06-2009:22:00
saurabh	87657	25-06-2009:22:00
disc	848853349	25-06-2010:22:00
maharishi	1475777225	25-06-2009:22:00
[illegible]	[illegible]	09:22:00
[illegible]	[illegible]	09:22:00
[illegible]	[illegible]	09:22:00
Node1	771344924	25-06-2009:22:00

Figure 10: A SPIG(CMMS) working scenario.

7. Conclusion

Grid is a useful and exciting way of establishing large scale distributed and sharing of data, resources and services. Service oriented grid and VO for service rendering are also of prime importance as any other grid solution. Providing large scale distributed services requires high level of authentication and authorization. Existing grid infrastructures rely mainly on PKI and GSI. The available authorization by use of PEP or PDP, Push-Pull models are not addressing all shortcomings in grid authorization namely lack of statefulness, fine grained policies and services, multi nodal policies etc. Though many researchers pointed out suggestions, a combined solution for any infrastructure is yet to come out. The CMMS suggests multilevel authorization, one at VO level and one at service level. Apart from this it introduces the concept of states to both user's, VO and service nodes. Although it suggests one way state-policy to service mapping. The two-way model is very similar and can be framed very easily. The proposed model requires certain technological changes in PKI and X.509. Mainly states needs to be monitored through MONITOR nodes in CA and X.509 certificates should contain state or policy or both together. SPIG is a prototype implementation to understand the essence of the whole proposed CMMS model in a very practical grid services scenario.

8. Acknowledgment

Our sincere thanks to all friends and colleagues who supported us at each respect and showed us our strengths and weaknesses. Lastly, heart-full thanks to Payal and Rimi.

9. References

[1] Grid security infrastructure. http://www.globus.org/security/overview.html.

[2] Sarbjeet Singh, and Seema Bawa, A Privacy, Trust and Policy based Authorization Framework for Services in Distributed Environments, International Journal of Computer Science Volume 2 Number 2, 2007.

[3] K. M. Chandy and L. Lamport. Distributed snapshots: Determining global states of distributed systems. ACM Transactions on Computer Systems, 3(1):63–75, Feb. 1985.

[4] Adam J. Lee, Marianne Winslett, Safety and Consistency in Policy Based Authorization, Department of Computer Science, University of Illinois at Urbana Champaign 2006.

[5] S. Farrell, J. Vollbrecht, P. Calhoun, L. Gommans, G. Gross, B. De Bruijn, C. De Laat, M. Holdrege, D. Spence. AAA Authorization Requirements. Request For Comments 2906, Network Working Group (2000). http://www.Ietf.Org/Rfc/Rfc2096.Txt.

[6] J. Vollbrecht, P. Calhoun, S. Farrell, L. Gommans, G. Gross, B. De Bruijn, C. De Laat, M. Holdrege, D. Spence. AAA Authorization Framework Request For Comments 2904, Network Working Group (2000). http://www.Ietf.Org/Rfc/Rfc2904.Txt.

[7] M. Lorch, D. Adams, D. Kafura, M. Koneni, A. Rathi, and S. Shah. The PRIMA System for Privilege Management, Authorization and Enforcement in Grid Environments. In Proceedings of the 4th Int. Workshop on Grid Computing - Grid 2003, Phoenix, AZ, USA, Nov. 2003.

[8] R. Alfieri, R. Cecchini, V. Ciaschini, L. dell'Agnello, A. Frohner, A. Gianoli, K. L˝orentey, and F. Spataro. VOMS: An Authorization System for Virtual Organizations. In proceedings of the 1st European Across Grids Conference, Santiago de Compostela, Feb. 2003.

[9] Laura Pearlman, Von Welch, Ian Foster, Carl Kesselman, Steven Tuecke. A Community Authorization Service for Group Collaboration. http://www.globus.org/alliance/publications/papers/CAS_2002_Revised.pdf

[10] S Farrell, R Housley. An Internet Attribute Certificate Profile for Authorization. Request For Comments 3281, Network Working Group (2002).

[11] M. R. Thompson, A. Essiari, K. Keahey, V, Welch, S. Lang, B. Liu, Fine-Grained Authorization for Job and Resource Management Using Akenti and the Globus Toolkit, CHEP 03, La Jolla, Mar 24-28, 2003.

[12] Bart Jacob, Michael Brown, Kentaro Fukui, Nihar Trivedi. Introduction to Grid Computing, ibm.com/redbooksibm.com/redbooks. IBM Corporation Dec 2005.

[13] Marty Humphrey, Member, IEEE, Mary R. Thompson, Member, IEEE, And Keith R. Jackson, Security for Grids, Proceedings Of The IEEE, Vol. 93, No. 3, March 2005.

[14] R. Gröper, Policy-based Authorization for Grid Data-Management, Master Thesis, University of Hannover, 2006.

[15] J. Basney, W. Nejdl, D. Olmedilla, V. Welch, and M. Winslett, "Negotiating trust on the grid," in Proc. 2nd Workshop on Semantics in P2P and Grid Computing, 2004. www.ncsa.illinois.edu/~jbasney/sempgrid.pdf.

[16] Guangsen Zhang, Manish Parashar, Dynamic Context-aware Access Control for Grid Applications, Proceedings of the Fourth International Workshop on Grid Computing (GRID'03), IEEE 2003.

[17] Ionut Constandache, Daniel Olmedilla, and Wolfgang Nejdl, Policy Based Dynamic Negotiation for Grid Services Authorization, Semantic Web Policy Workshop, ISWC'05 Alway, Ireland, 7th November 2005.

Criminal and Civil Identification with DNA Databases Using Bayesian Networks

Marina Andrade marina.andrade@iscte.pt
Department of Quantitative Methods
ISCTE – Lisbon University Institute
Lisbon, 1649-026, Portugal

Manuel Alberto M. Ferreira manuel.ferreira@iscte.pt
Department of Quantitative Methods
ISCTE – Lisbon University Institute
Lisbon, 1649-026, Portugal

Abstract

Forensic identification problems are examples in which the study of DNA profiles is a common approach. Here we present some problems and develop their treatment putting the focus in the use of Object-Oriented Bayesian Networks - OOBN. The use of DNA databases, which began in 1995 in England, has created new challenges about its use. In Portugal the legislation for the construction of a genetic database was defined in 2008. With this it is important to determine how to use it in an appropriate way.
For a crime that has been committed, forensic laboratories identify genetic characteristics in order to connect one or more individuals to it. Apart the laboratories results it is a matter of great importance to quantify the information obtained, i.e., to know how to evaluate and interpret the results obtained providing support to the judicial system. Other forensic identification problems are body identification; whether the identification of a body (or more than one) found, together with the information of missing persons belonging to one or more known families, for which there may be information of family members who claimed the disappearance. In this work we intend to discuss how to use the database; the hypotheses of interest and the database use to determine the likelihood ratios, i.e., how to evaluate the evidence for different situations.

Keywords: Bayesian networks, DNA profiles, identification problems.

1. INTRODUCTION

The use of networks transporting probabilities began with the geneticist Sewall Wright in the beginning of the 20th century (1921). Since then their use had different forms in several areas like social sciences and economy – in which the used models are, in general, linear named Path Diagrams or Structural Equations Models (SEM), and in artificial intelligence – usually non-linear models named Bayesian networks also called Probabilistic Expert Systems (PES), [11],[14].

Bayesian networks are graphical structures for representing the probabilistic relationships among a large number of variables and for doing probabilistic inference with those variables, [13]. Before we approach the use of Bayesian networks to our interest problems we briefly discuss some aspects of PES in connection with uncertainty problems in section 2.

In section 3 after presenting some possible forensic identification problems the creation and use of DNA profile databases in some European countries is discussed putting the focus in the entry criteria and the differences observed. How to approach evaluate and interpret the results, whether it is a criminal or a civil identification problem is presented in section 4. Thus, it is important to describe the Portuguese law establishing the principles to maintain a DNA database file for civil and criminal identification purposes. The study of a criminal identification problem considering one single perpetrator, and a civil identification problem with one volunteer, for two different situations, are considered.

2. EXPERT SYSTEMS

Expert systems are attempts to crystallize and codify the knowledge and skills of one or more experts into a tool that can be used by non-specialists, [14]. An expert system can be decomposed as follows:

Expert system = knowledge base + Inference engine.

The first term on the right-hand side of the equation, knowledge base, refers the specific knowledge domain of the problem. The inference engine is given by a set of algorithms, which process the codification of the knowledge base jointly with any specific information known for the application in study.

Usually it is presented in a software program, as the one we are going to show hereafter, but such is not an imperative rule. Each of those parts is important for the inferences, but knowledge base is crucial. The inferences obtained depend naturally on the quality of the knowledge base, of course in association with a sophisticated inference engine. The better those parts are the best results we can get.

A PES is a representation of a complex probability structure by means of a directed acyclic graph, having a node for each variable, and directed links describing probabilistic causal relationships between variables, [1]. Bayesian approach is the adequate for making inferences in probabilistic expert systems.

2.1 Bayesian networks

Bayesian networks are graphical representations expressing qualitative relationships of dependence and independence between variables. A Bayesian network is a directed acyclic graph $\mathcal{G}$ (DAG) having a set of V vertices or nodes and directed arrows. Each node $\upsilon \epsilon V$ represents a random variable X_υ with a set of possible values or states. The arrows connecting the nodes describe conditional probability dependencies between the variables.

The set of parents, $pa(\upsilon)$, of a node υ comprises all the nodes in the graph with arrows ending in υ. The probability structure is completed by specifying the conditional probability distributions for each random variable X_υ and each possible configuration of variables associated with its parent nodes $x_{pa(\upsilon)}$. The conditional distribution of X_υ is expressed given $X_{pa(\upsilon)} = x_{pa(\upsilon)}$. The joint distribution is $p(x) = \Pi_{\upsilon \epsilon V} p(x_\upsilon | x_{pa(\upsilon)})$. There are algorithms to transform the network into a new graphical representation, named junction tree of cliques, so that the conditional probability $p(x_\upsilon | x_A)$ can be efficiently computed, for all $\upsilon \epsilon V$, any set of nodes $A \subseteq V$, and any configuration x_A of the nodes X_A. The nodes in the conditioning set A are generally nodes of observation and input of evidence $X_A = x_A$, or they may specify hypotheses being assumed.

Software such as Hugin[1] can be used to build the Bayesian network through the graph $\mathcal{G}$. That can be done by specifying the graph nodes, their space of states and the conditional probabilities $p(x_v| x_{pa(v)})$. In the compiling process the software will construct its internal junction tree representation. Then, by entering the evidence $X_A= x_A$ at the nodes in A, and requesting its propagation to the remaining nodes in the network, the conditional probabilities $p(x_v| x_A)$ are obtained. Version 6.4 of Hugin and upgrades allow the graphical use of OOBN.

OOBN are one example of the general class of Bayesian networks. An instance or object is a regular network possessing input and output nodes as well as ordinary internal nodes. The interface nodes have grey fringes, with the input nodes exhibiting a dotted line and the output nodes a solid line. The instances of a given class have identical conditional probability tables for non-input nodes. The objects are connected by directed links from output to input nodes. The links represent identification of nodes. We use bold face to refer the object classes and math mode to refer the nodes. The modular flexibility structure of the OOBN is of great advantage in complex cases

3. Forensic identification problems

The use of DNA profiles in forensic identification problems has become, in the last years, an almost regular procedure in many and different situations. Among those are: 1) disputed paternity problems, in which it is necessary to determine if the putative father of a child is or is not the true father; 2) criminal cases as if a certain individual Y was the origin of a stain found in the scene of a crime; or in more complex cases to determine if an individual or more did contribute to a mixture trace found; 3) civil identification problems, i.e., the case of a body identification, together with the information of a missing person belonging to a known family, or the identification of more than one body resultant of a disaster or an attempt. And even immigration cases in which it is important to establish family relations.

Here the focus is to approach the civil and criminal identification problems. The establishment and use of DNA database files for a great number of European countries worked as a motivation to study in more detail the mentioned problems and the use of these database files identification.

The use of a DNA profile database may allow delinquents' identification and/or the connection of criminal conducts and the respective individuals, the exclusion of innocents as well as the recognition and civil identification of missing people. A genetic profile database may be an important help in forensic investigation, particularly in crimes of repetitive tendency, when DNA samples of condemned individuals are collected. In the context of the civil identification it may be very useful when unidentified corpses appear and may be identified by comparison of their DNA profiles family volunteer's profiles.

3.1 DNA database files

The discovery of biological fingerprints in 1984 opened new perspectives in forensic identification area. Since then the technical advances and results obtained have allowed studying and solving increasingly complex problems. Almost twenty five years after Alec Jeffreys' team discovery, Portugal established the legislation for the construction of a genetic database, law nº5/2008.

The advances in DNA technology and knowledge opened new perspectives for civil and criminal investigation. Apart from the ethical and legal questions that are in the domain of the legal system, we should draw some attention to the experience and knowledge acquired by those countries that already have their own databases operating and try to learn from them ways to improve on how to operate with the Portuguese database, [5].

[1] http://www.hugin.com - OOBN a resource available in the Hugin 6.4 software.

Country	Year	Entry criteria for suspects (convicted offenders)
England	1995	Any recordable offence
Austria	1997	Any recordable offence
Croatia/Switzerland	2000	Any recordable offence
Germany	1998	> 1 year in prison (after court decision)
Finland	1999	> 1 year in prison
Denmark	2000	> 1.5 years in prison
Norway	1999	Many serious offences (after court decision)
Hungary	2003	5 years in prison
Sweden	2000	No suspects entered (> 2 years convicted)
Belgium	2002	No suspects entered (after court decision)
Netherlands	1997	No suspects entered (> 4 years convicted)
France	2001	No suspects entered (serious offences, voluntary samples only)
Spain	1998	Phoenix program – civil database for civil ident. vol. donations
Portugal	2008	Vol.,“problem samples”, “reference samples” (≥ 3 years convicted)
Italy	-	Law in preparation

TABLE 1: National DNA databases in Europe.

There has been a considerable discussion about the individuals to include in a DNA profiles database, specially with different results in countries with different legal systems. The main differences are linked to the emphasis given by the countries: to the individuals or the social order.

In the table above (TABLE 1) it is possible to observe significant differences between the European countries in what concerns criteria to enter a person into a database (while a suspect or only after conviction, different types of conviction). The criteria to remove records and the number of entries in the database also have important differences all over Europe.

As we have seen there are clear differences between countries more on the north and more on the south of Europe, which is essentially due the different perspective valuing more the social order or the individual itself.

4. Criminal and civil identification using DNA profile databases

Let us consider a criminal case in which a DNA profile has been recovered from a crime scene, and it is admitted to be left by the culprit (only one perpetrator); and a civil identification problem with one volunteer giving his/her genetic information.

The Portuguese law n°5/2008 establishes the principles for creating and maintaining a database of DNA profiles for identification purposes, and regulates the collection, processing and conservation of samples of human cells, their analysis and collection of DNA profiles, the methodology for comparison of DNA profiles taken from the samples, and the processing and storage of information in a computer file.

Here it is assumed that the database is composed of a file containing information of samples from convicted offenders with 3 years of imprisonment or more - α; a file containing the information of samples of volunteers - β; a file containing information on the "problem samples" or "reference samples" from corpses, or parts of corpses, or thing or place where the authorities collect samples - γ.

4.1 Criminal identification – one single perpetrator

For a crime that has been committed, forensic laboratories identify genetic characteristics in order to connect one or more individuals to the crime. Apart from the laboratories results it is a matter of great importance to quantify the information obtained, i.e., to know how to evaluate and interpret the results obtained providing a support to the judicial decision. Here it is assumed a DNA trace found at a scene of a crime left by only one perpetrator.

“The experience with some databases seams to indicate that before the commitment of a serious crime, some suspects have already been involved in minor offences. This fact associated to the repetitive motif of some serious crimes can support the importance of DNA databases not only to the criminal investigation but also to the prevention of crime, mainly if there are large entry criteria.”, [4].

Some notation:
Let C_c be the genetic characteristic (DNA profile) found at the crime scene, and C_s the suspect’s genetic characteristic. The evidence is $E=(C_c,C_s)$, the DNA typing results for the crime sample and the suspect. Our interest is to discuss how to evaluate the hypotheses and the *odds* ratio. In court the hypotheses are: *H_P*: The suspect *(s)* left the crime stain *vs H_D*: Some other person left the crime stain.

There is a match between the crime scene profile and the suspect’s profile. The court wants to compare the two preceding hypotheses. It is important to discuss the presentation of the evidence in court and how to evaluate the hypotheses of interest. The *posterior odds* are:

$$\frac{P(H_P \mid E, s\in\alpha)}{P(H_D \mid E, s\in\alpha)} = \frac{P(E \mid H_P, s\in\alpha)}{P(E \mid H_D, s\in\alpha)}\frac{P(H_P \mid s\in\alpha)}{P(H_D \mid s\in\alpha)}$$

$$= \underbrace{\frac{P(C_c, C_s \mid H_P, s\in\alpha)}{P(C_c, C_s \mid H_D, s\in\alpha)}}_{LR}\frac{P(H_P \mid s\in\alpha)}{P(H_D \mid s\in\alpha)}$$

The likelihood ratio, LR, takes the form:

$$LR = \frac{P(C_c \mid C_s, H_P, s\in\alpha)}{P(C_c \mid C_s, H_D, s\in\alpha)}\frac{P(C_s \mid H_P, s\in\alpha)}{P(C_s \mid H_D, s\in\alpha)}$$

Whether or not the suspect left the crime sample that does not provide any information to our uncertainty about his/her genetic characteristic or genotype, i.e., $P(C_s|H_P, s\in\alpha) = P(C_s|H_D, s\in\alpha)$. Therefore

$$LR = \frac{P(C_c \mid C_s, H_P, s\in\alpha)}{P(C_c \mid C_s, H_D, s\in\alpha)}.$$

In a case with a trace of a single perpetrator, for each marker, one can use the object-oriented Bayesian (OOBN) network shown in Figure 1. Each node (instance) in the network represents itself a Bayesian network. Nodes **spg** and **smg** are all of class founder, a network with only one node which states are the alleles in the problem and the respective frequencies in the population, and represent the suspect's paternal and maternal inheritance. Node **sgt** and **cgt** are of class genotype. The genotype of an individual is an unordered pair of alleles inherited from paternal (*pg*) and maternal (*mg*) genes, here represented by *gtmin:=min*{*pg*, *mg*} and *gtmax:=max*{*pg*, *mg*}, where *pg* and *mg* are input nodes identical to the *gene* node of *founder***.** The nodes **cmg** and **cpg** specify whether the correspondent allele is or is not from the suspect. If **s=c?** has true for value then the true perpetrator's allele will be identical with the suspect's allele, otherwise the true perpetrator's allele is chosen randomly from another man in the population. The single node **s=c?** represent the binary query 'Is the suspect the perpetrator?'

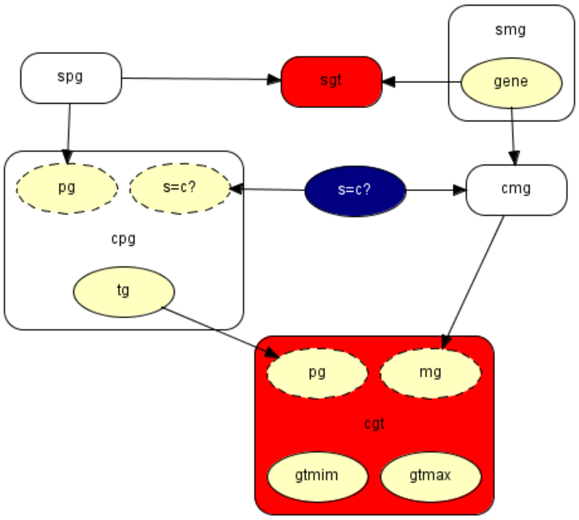


FIGURE 1: Network for a criminal case with a single perpetrator.

Here, if the suspect, s, is in the file of convicted offenders then $P(H_P \mid s\in\alpha) > P(H_D \mid s\in\alpha)$, otherwise those probabilities may be assumed equal. Therefore, the posterior odds are:

If s is in the file of convicted offenders Otherwise

$$\frac{1}{P(C_c \mid C_s, H_D, s\in\alpha)} \underbrace{\frac{P(H_P \mid s\in\alpha)}{P(H_D \mid s\in\alpha)}}_{>1} > \frac{1}{P(C_c \mid C_s, H_D, s\in\alpha)}.$$

We need to assess $P(H_P \mid s \in \alpha)$ and $P(H_D \mid s \in \alpha)$. The posterior odds computation is in the judges' domain, we can only explain how to do it and how to interpret the evidence, i.e., how to use it as a decision support element.

4.2 Civil identification – one missing person and one volunteer

A missing individual is reported. A body is found. The hypotheses are: *H_P*: The body found is the body of the claimed individual *X vs H_D*: The body found is any other individual's, not the claimed individual *X*. People want their relatives alive and reject the hypothesis that states their lost relative is dead. A volunteer supplies genetic material to be used in the test of a partial match. The evidence is $E = (C_{Bf}, C_{vol})$ - the genetic characteristics of the body found, C_{BF}, and the volunteer, C_{vol}. The *posterior odds* is

$$\frac{P(H_P \mid E, vol \in \beta, \gamma)}{P(H_D \mid E, vol \in \beta, \gamma)} = \frac{P(E \mid H_P, vol \in \beta, \gamma)}{P(E \mid H_D, vol \in \beta, \gamma)} \frac{P(H_P \mid vol \in \beta, \gamma)}{P(H_D \mid vol \in \beta, \gamma)}$$

One may assume $P(H_P \mid vol \in \beta, \gamma) = P(H_D \mid vol \in \beta, \gamma)$ then

$$\frac{P(H_P \mid E, vol \in \beta, \gamma)}{P(H_D \mid E, vol \in \beta, \gamma)} = \underbrace{\frac{P(C_{BF}, C_{vol} \mid H_P, vol \in \beta, \gamma)}{P(C_{BF}, C_{vol} \mid H_D, vol \in \beta, \gamma)}}_{LR}$$

$$\frac{P(H_P \mid E, vol \in \beta, \gamma)}{P(H_D \mid E, vol \in \beta, \gamma)} = \frac{P(C_{BF} \mid C_{vol}, H_P, vol \in \beta, \gamma)}{P(C_{BF}, \mid C_{vol}, H_D, vol \in \beta, \gamma)} \underbrace{\frac{P(C_{vol} \mid H_P, vol \in \beta, \gamma)}{P(C_{vol} \mid H_D, vol \in \beta, \gamma)}}_{=1}.$$

The conditioning does not include the information of the body found. Whether or not the volunteer is related with the individual whose body was found that does not provide information to our uncertainty about his genotype.

Depending on the volunteer and the claimed individual family relation we only may observe a partial match with one volunteer. Apart form that, it is important to check if there is a match between C_{BF} and any of the "problem samples" in γ. Assuming no match of C_{BF} and any sample in γ, the likelihood ratio may be written as:

$$\frac{P(C_{BF} \mid C_{vol}, H_P)}{P(C_{BF} \mid H_D)}.$$

In a case having only one volunteer the likelihood ratio can be computed using a Bayesian network. Suppose we have a missing individual, an elder person and a son or a daughter claiming the disappearance and who gives a genetic profile voluntarily. The likelihood ratio can be computed using the following network:

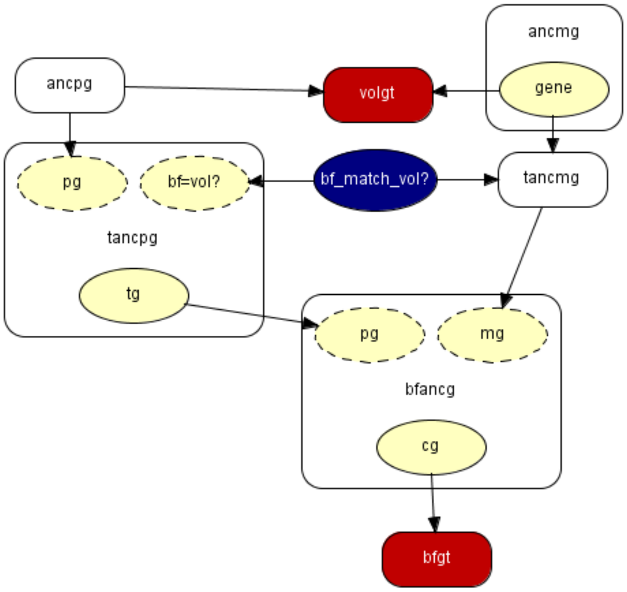


FIGURE 2: Network for civil identification with one volunteer, paternal or maternal relationship case.

As above nodes **ancpg** and **ancmg** are all of class founder, a network with only one node which states are the alleles in the problem and the respective frequencies in the population, and represent the volunteer's ancient paternal and maternal inheritance. Node **volgt** and **bfgt** are of class genotype, the volunteer and body found genotypes.

The nodes **tancmg** and **tancpg** specify whether the correspondent allele is or is not from the volunter. If **bf_match_vol?** has true for value then the volunteer's allele will be identical with the volunteer's allele. The node **bfancg** defines the Mendel inheritance in which the allele of the individual whose body was found is chosen at random from the ancient's paternal and maternal gene.

For a case with only one volunteer, but now for example a brother or a sister of an individual who is missing and is being searched, the likelihood ratio may be computed using the following Bayesian network:

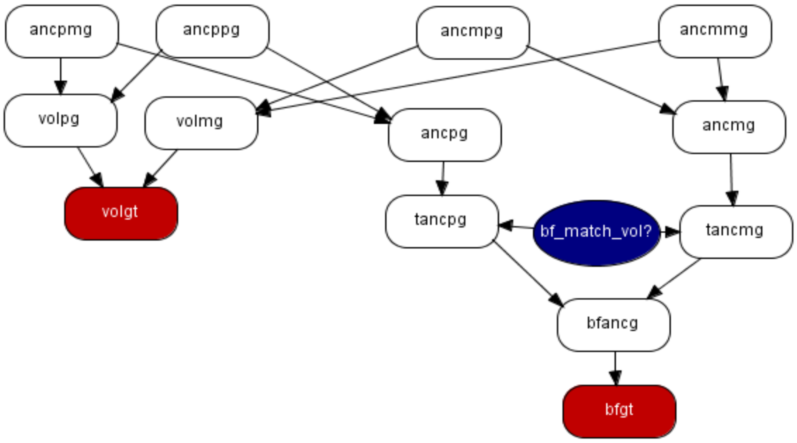


FIGURE 3: for civil identification with one volunteer, brother or sister relationship case.

Nodes **ancpmg** and **ancppg** are all of class founder and represent the volunteer's ancient genes of paternal and maternal inheritance. That inheritance will pass to nodes **volpg** and **volmg**, which are going to form volunteer genotype. Node **volgt** and **bfgt** are of class genotype, the volunteer and body found genotypes. The remaining nodes are the same as the ones presented in the previous problem.

5. CONCLUSION & FUTURE WORK

To connect an individual with a crime on the basis of a profile match may be dangerous because the database may contain undetected errors. In order to avoid misclassification with DNA from the database it is important to admit, at least, a second and independent analysis.

After computing the likelihood, whether it is a criminal case or a civil identification case, it is possible to compute the posterior odds, i.e., multiplying the likelihood ratio and the prior odds, in order to perform a comparative evaluation between the prosecution and the defense hypotheses.

The database file α is a subset of the population set P, $\alpha \subset P$. If the size of the database file is small, then one may only have a small fraction of the possible offenders. Therefore, it is important to take that into account. This topic should be considered in future work.

Whether it is criminal or civil identification in many situations the evidence may have more than one individual involved. In future work that must be considered. Also, in civil identification problems an important issue is to study how to compute the likelihood ratios when there is a match or a partial match between the genetic characteristic of the individual whose body was found and the file of "problem samples" and "reference samples", γ.

6. REFERENCES

1. A. P. Dawid, J. Mortera, V. L. Pascali, D. W. Van Boxel. *"Probabilistic expert systems for forensic inference from genetic markers"*. Scandinavian Journal of Statistics, 29:577-595, 2002

2. B. P. Battula, K. Rani , S. Prasad , T. Sudha. *"Techniques in Computer Forensics: A Recovery Perspective"*. International Journal of Security, Volume 3, Issue 2:27-35, 2009

3. David J. Balding. *"The DNA database controversy"*. Biometrics, 58(1):241-244, 2002

4. F. Corte-Real. *"Forensic DNA databases"*. Forensic Science International, 146s:s143-s144, 2004

5. G. Skinner. *"Multi-Dimensional Privacy Protection for Digital Collaborations"*. International Journal of Security, Volume 1, Issue 1:22-31, 2007

6. I. Evett and B. S. Weir. *"Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists"*, Sinauer Associates, Inc. (1998)

7. M. Andrade, M. A. M. Ferreira. "*Bayesian networks in forensic identification problems*". Journal of Applied Mathematics. Volume 2, number 3, 13-30, 2009

8. M. Andrade, M. A. M. Ferreira, J. A. Filipe. *"Evidence evaluation in DNA mixture traces"*. Journal of Mathematics and Allied Fields (Scientific Journals International-Published online). Volume 2, issue 2, 2008

9. M. Andrade, M. A. M. Ferreira, J. A. Filipe., M. Coelho. *"Paternity dispute: is it important to be conservative?"*. Aplimat – Journal of Applied Mathematics. Volume 1, number 2, 2008

10. M. Guillén, M. V. Lareu, C. Pestoni, A. Salas and A. Carrecedo. *"Ethical-legal problems of DNA databases in criminal investigation"*. Journal of Medical Ethics, 26:266-271, 2000

11. M. N. Anyanwu, S. Shiva. *"Comparative Analysis of Serial Decision Tree Classification Algorithms"*. International Journal of Computer Science and Security, Volume 3, Issue 3:230-240, 2009

12. P. Martin. *"National DNA databases – practice and practability. A forum for discussion"*. In International Congress Series 1261, 1-8

13. R. E. Neapolitan. *"Learning Bayesian networks"* , Pearson Prentice Hall, (2004)

14. R. G. Cowell, A. P. Dawid, S. L. Lauritzen, D. J. Spiegelhalter. *"Probabilistic Expert Systems"*, Springer, New York, (1999).

Performance Evaluation and Comparison of On Demand Multicast Reactive Routing Protocols under Black Hole Attack in MANET

E. A. Mary Anita anitareginald@yahoo.co.in
Research Scholar
Anna University
Chennai, India

V. Vasudevan drvvmca@yahoo.com
Senior Professor and Head / IT
A. K. College of Engineering
Virudunagar, India

Abstract

One main challenge in the design of routing protocols is their vulnerability to security attacks. This is mainly due to the wireless and dynamic nature of ad hoc networks. A black hole attack is a severe attack that can be easily employed against routing in mobile ad hoc networks. In this attack a malicious node advertises itself as having the shortest path to the node whose packets it wants to intercept thereby exploiting the proper functioning of the protocol. In this paper the performance of multicast on demand routing protocols such as Multicast Ad-hoc On Demand Distance Vector (MAODV) protocol and On Demand Multicast Routing Protocol (ODMRP) are evaluated and analyzed under black hole attack under different scenarios in terms of the performance metrics such as packet delivery ratio and end to end delay. The evaluation is done with network simulator NS-2. Simulation results indicate that both the protocols suffer a significant reduction in packet delivery ratio in the presence of black hole attackers but the impact is more in MAODV when compared to ODMRP due to the presence of alternate data delivery paths in ODMRP.

Keywords: MANET, Black hole, MAODV, ODMRP, Packet Delivery Ratio, End to End Delay.

1. INTRODUCTION

Security in wireless ad-hoc networks is a complex issue. This complexity is due to various factors like insecure wireless communication links, absence of a fixed infrastructure, node mobility and resource constraints [1]. Nodes are more vulnerable to security attacks in mobile ad-hoc networks than in traditional networks with a fixed infrastructure. The security issues of Mobile Ad-hoc Networks (MANETs) are more challenging in a multicasting environment with multiple senders and receivers. There are different kinds of attacks by malicious nodes that can harm a network and make it unreliable for communication. These attacks can be classified as active and passive attacks [2]. A passive attack is one in which the information is snooped by an intruder without disrupting the network activity. An active attack disrupts the normal operation of a network by modifying the packets in the network. Active attacks can be further classified as internal and external attacks. External attacks are carried out by nodes that do not form part of the network. Internal attacks are from compromised nodes that were once legitimate part of the network.

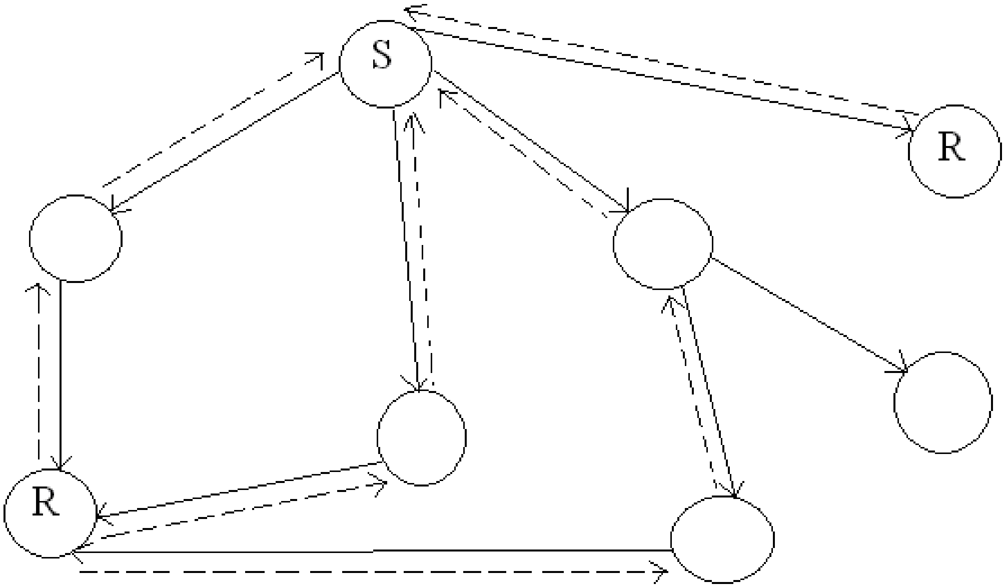
A black hole attack is one in which a malicious node advertises itself as having the shortest path to a destination in a network. This can cause Denial of Service (DoS) by dropping the received packets.

The rest of the paper is organized as follows. The next section gives an overview of MAODV and ODMRP. Section III discusses about black hole attack. In section IV the results of simulation experiments that show the impact of black hole attack on the performance of MAODV and ODMRP under different scenarios are discussed. Finally section V summarizes the conclusion

2. OVERVIEW OF ROUTING PROTOCOLS

2.1 Overview of MAODV

MAODV is a multicast routing protocol for ad-hoc networks. It is an extension of AODV. As nodes join the group, a tree is created. This tree connects the group members and many routers which are not group members but exist in the tree to connect the group members.



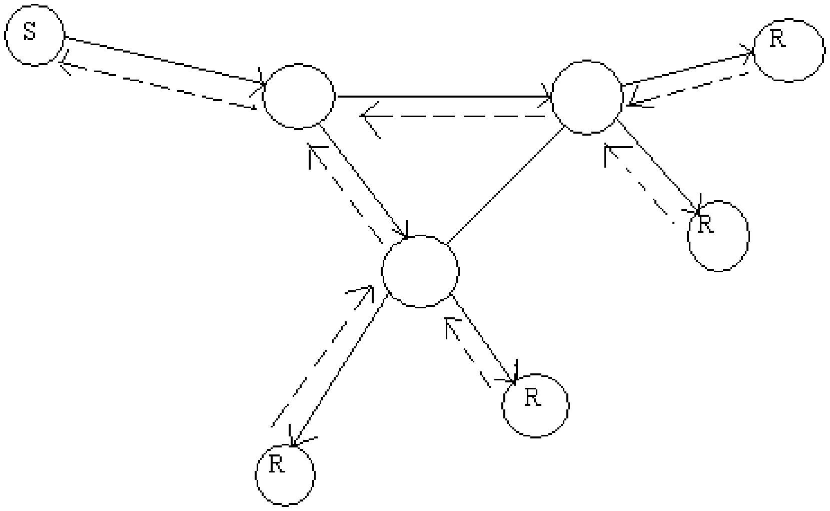
R – Multicast Receiver ________RREQ _ _ _ _ _ RREP

FIGURE 1: Route Request procedure of MAODV

Multicast group membership is dynamic. The group members and routers are all members of the tree. Every multicast group is identified by a unique address and group sequence number for tracing the freshness of the group condition [3]. When a node wants to find a route to a group or join a group it broadcasts a RREQ message. Any node with a fresh enough route to the multicast group may respond to this request message with a RREP message. If a node wants to become a member of the group that does not exist, then this node becomes the leader of that group and is responsible for maintaining the group. Group Hello messages are broadcasted periodically to check for connectivity of tree structure [4]. This results in increased overhead in maintaining route

2.2 Overview of ODMRP

ODMRP is a mesh based multicast routing protocol that uses the concept of forwarding group. Only a subset of nodes forwards the multicast packets on shortest paths between member pairs to build a forwarding mesh for each multicast group [5].



O – Mobile node S – Multicast Source R – Multicast Receiver

______ JREQ _ _ _ _ _ JREP

FIGURE 2: On demand route and mesh creation

In ODMRP, group membership and multicast routes are established and updated by the source on demand. When a multicast source has packets to send, it initiates a route discovery process. A JOIN REQUEST packet is periodically broadcast to the entire network. Any intermediate node that receives a non- duplicate JREQ packet stores the upstream node ID and rebroadcasts the packet. Finally when this packet reaches the destination, the receiver creates a JOIN REPLY and broadcasts it to its neighbors. Every node receiving the JREP checks to see if the next node id in JREP matches its own. If there is a match, it is a part of the forwarding group, sets its FG_FLAG and broadcasts its JREP built upon matched entries. This JREP is thus propagated by each forwarding group member until it reaches the source via a shortest path. Thus routes from sources to receivers build a mesh of nodes called forwarding group.

The forwarding group is a set of nodes that forward the multicast packets. It supports shortest paths between any member pairs. All nodes inside the bubble (multicast members and forwarding group nodes) forward multicast data packets [6]. A multicast receiver can also be a forwarding group node if it is on the path between a multicast source and another receiver. The mesh provides richer connectivity among multicast members compared to trees.
After the route establishment and route construction process, a multicast source can transmit packets to receivers via selected routes and forwarding groups. A data packet is forwarded by a node only if it is not a duplicate one and the setting of the FG_Flag for the multicast group has not expired. This procedure minimizes traffic overhead and prevents sending packets through stale routes.

In ODMRP, no explicit control packets need to be sent to join or leave the group. A multicast source can leave the group by just stop sending JREQ packets when it does not have any data to be sent to the group. If a receiver no longer wants to receive data from a particular group, it removes the corresponding entries from its member table and does not transmit the JOINTABLE for that group.

3. BLACK HOLE ATTACK

Routing protocols are exposed to a variety of attacks. Black hole attack is one such attack in which a malicious node makes use of the vulnerabilities of the route discovery packets of the routing protocol to advertise itself as having the shortest path to the node whose packets it wants to intercept [7]. This attack aims at modifying the routing protocol so that traffic flows through a specific node controlled by the attacker. During the route discovery process, the source node sends route discovery packets to the intermediate nodes to find fresh path to the intended destination. Malicious nodes respond immediately to the source node as these nodes do not refer the routing table.

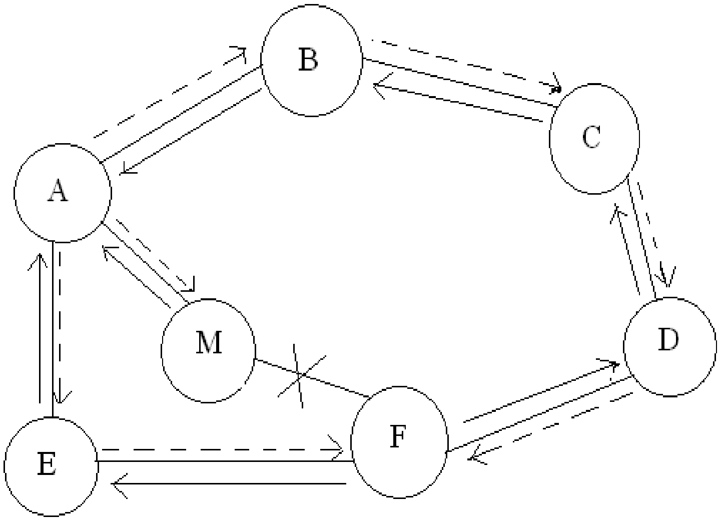
The source node assumes that the route discovery process is complete, ignores other route reply messages from other nodes and selects the path through the malicious node to route the data packets. The malicious node does this by assigning a high sequence number to the reply packet [8]. The attacker now drops the received messages instead of relaying them as the protocol requires.

3.1 Security in MAODV and ODMRP

ODMRP and MAODV are important on demand multicasting routing protocols that create routes only when desired by the source node. These protocols do not include any provisions for security and hence they are susceptible to attacks .When a node requires a route to a destination it initiates a route discovery process within the network. Any malicious node can interrupt this route discovery process by claiming to have the shortest route to the destination thereby attracting more traffic towards it [9].

For example, source A wants to send packets to destination D, in figure 3; source A initiates the route discovery process. Let M be the malicious node which has no fresh route to destination D. M claims to have the route to destination and sends route reply/join reply (RREP/JREP) packet to S. The reply from the malicious node reaches the source node earlier than the reply from the legitimate node, as the malicious node does not have to check its routing table like the other legitimate nodes [14].

The source chooses the path provided by the malicious node and the data packets are dropped [10].The malicious node forms a black hole in the network and this problem is called black hole problem.



A-Source node D-Destination node M-Malicious node

- - - -RREQ/JREQ ____ RREP/JREP

FIGURE 3: Black hole attack

4. SIMULATION

In this section, the simulation environment and the simulation results are discussed. Simulation is done using the network simulator NS-2.

4.1 Simulation Metrics

The metrics used in evaluating the performance are:

4.1.1 Packet Delivery Ratio: It is the ratio of the number of data packets delivered to the destinations to the number of data packets generated by the sources. This evaluates the ability of the protocol to deliver data packets to the destination in the presence of malicious nodes [11].

4.1.2 Average End-to-End Delay: This is the average delay between the sending of packets by the source and its receipt by the receiver [12]. This includes all possible delays caused during data acquisition, route discovery, queuing, processing at intermediate nodes, retransmission delays, propagation time, etc. It is measured in milliseconds.

4.2 Simulation Profile

The simulation settings are as follows. The network consists of 50 nodes placed randomly within an area of 1000m x 1000 m. Each node moves randomly and has a transmission range of 250m. The random way point model is used as the mobility model. In this model, a node selects a random destination and moves towards that destination at a speed between the pre-defined maximum and minimum speed. The minimum speed for the simulations is 0 m/s while the maximum speed is 50 m/s. The channel capacity is set to 2Mbps and the packet size is 512 bytes. The CBR traffic is generated with a rate of 4 packets per second. The simulation time is 900 seconds. The simulations were carried out with 0, 2 and 5 attackers for different number of receivers. The malicious nodes were selected randomly.

4.3 Discussion of results

Figure 4 shows the variation of packet delivery ratio (PDR) with mobility when the multicast group consists of 1 sender and 20 receivers with no attackers.

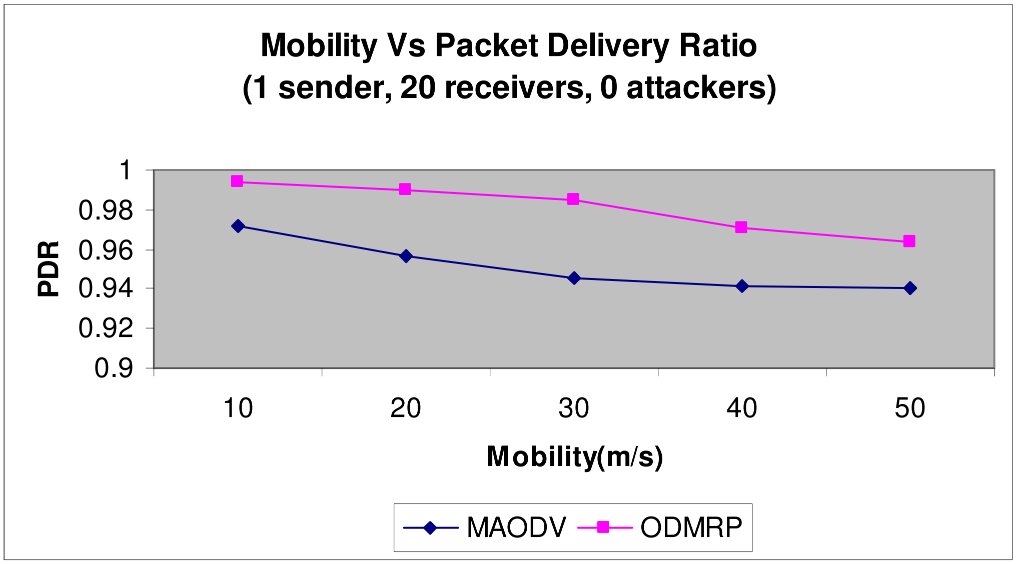


FIGURE 4: Variation of PDR with mobility in the absence of attackers

It is seen that the PDR decreases with increased mobility. Also the PDR of MAODV is less than the PDR of ODMRP by around 2 to 10%. This may be attributed to the fact that more alternate routing paths are available in ODMRP. The mesh structure in ODMRP provides multiple paths spanning all multicast group members and these paths become available in case of any failure in the primary path.

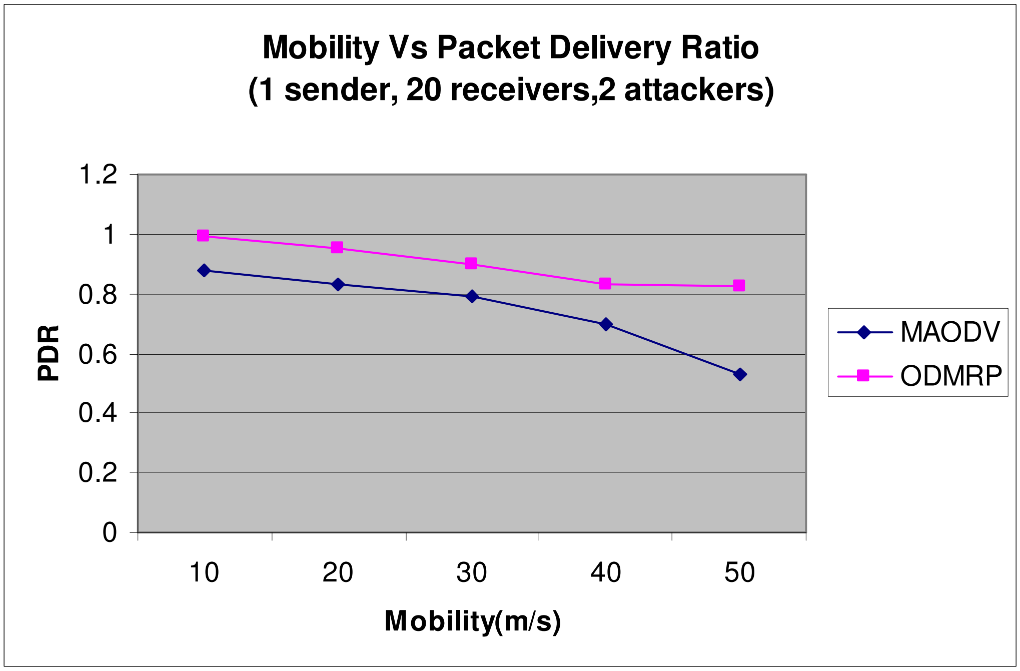


FIGURE 5: Variation of PDR with mobility in the presence of 2 attackers

When there are 2 numbers of attackers, the PDR reduces to about 1 to 4% for ODMRP and the reduction is around 5% to 20% in MAODV as shown in figure 5. This loss is partially due to black hole nodes dropping the packets and partially due to congestion in the network over the paths towards the black hole nodes.

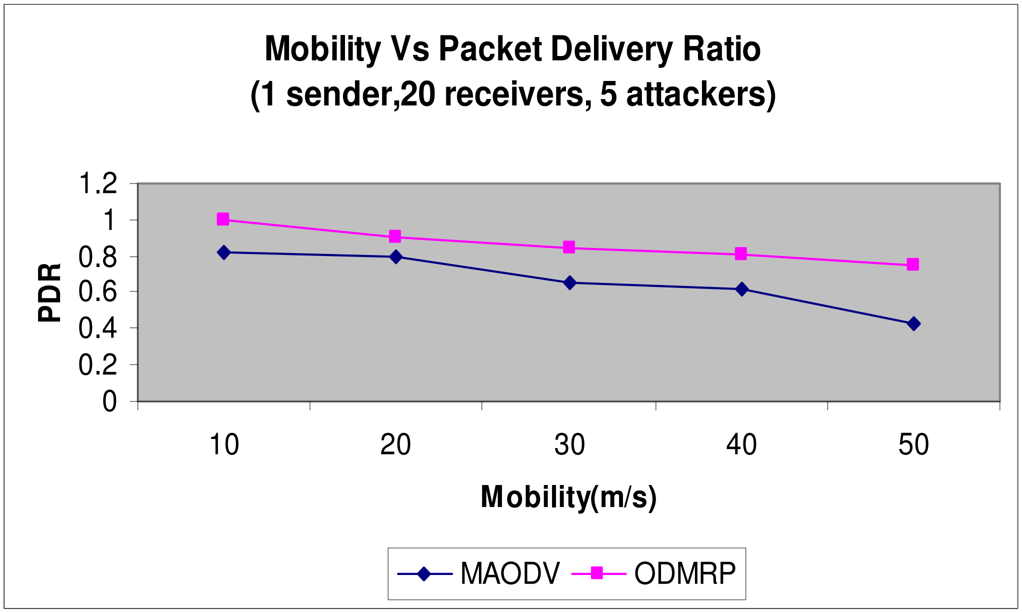


FIGURE 6: Variation of PDR with mobility in the presence of 5 attackers

When the number of attackers is increased to 5, the PDR further drops by around 5 for ODMRP and 20% for MAODV. Higher the number of attackers, higher the reduction in PDR. This is shown in figure 6.

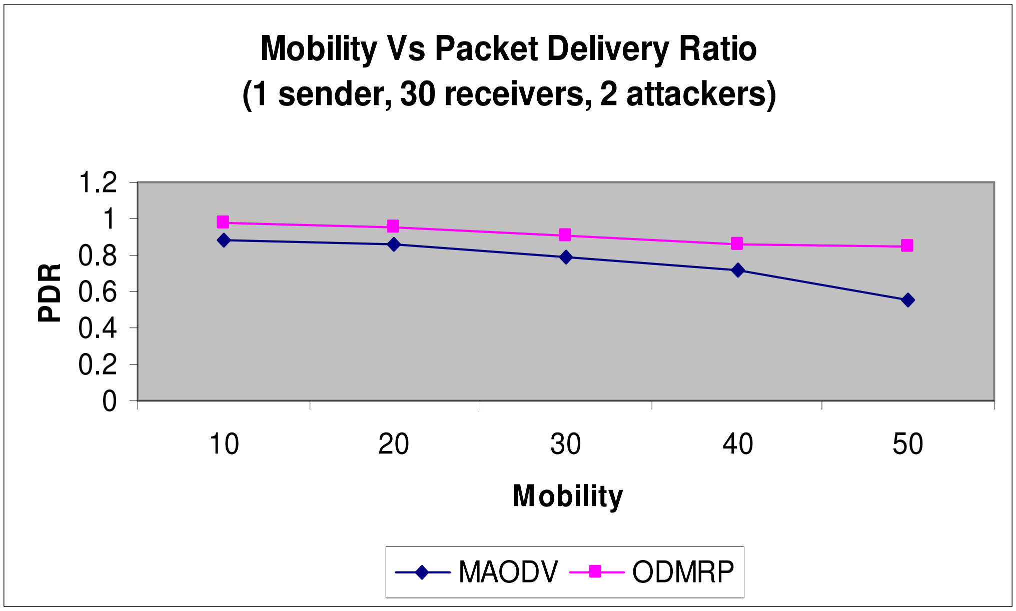


FIGURE 7: Variation of PDR with mobility in the presence of 30 receivers and 2 attackers

Figure 7 shows the variation of PDR with mobility with an increased group size of 30 receivers. It is seen that though the PDR reduces in the presence of attackers, a large group is able to withstand the attack to a reasonable extent when compared to a smaller group which is easily susceptible to attacks.

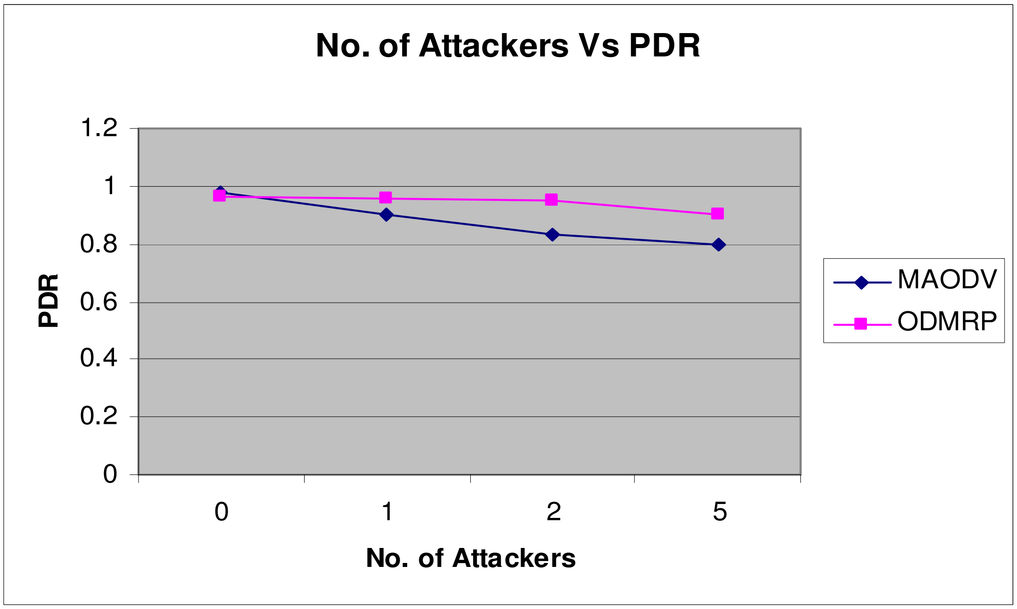


FIGURE 8: Variation of PDR with attackers

Figure 8 shows the variation of PDR for different number of attackers. It is seen that the packet delivery ratio reduces in the presence of attackers and the effect of the attack is more in MAODV when compared to ODMRP. This is due to the presence of alternate paths available in ODMRP. Since the mesh becomes denser with the growth if the members, more redundant routes are formed thereby improving the performance. So even if a packet gets dropped in one path due to the presence of black hole nodes, there is a chance for the duplicate copy of the packet to reach the destination through alternate paths free from malicious nodes [13].

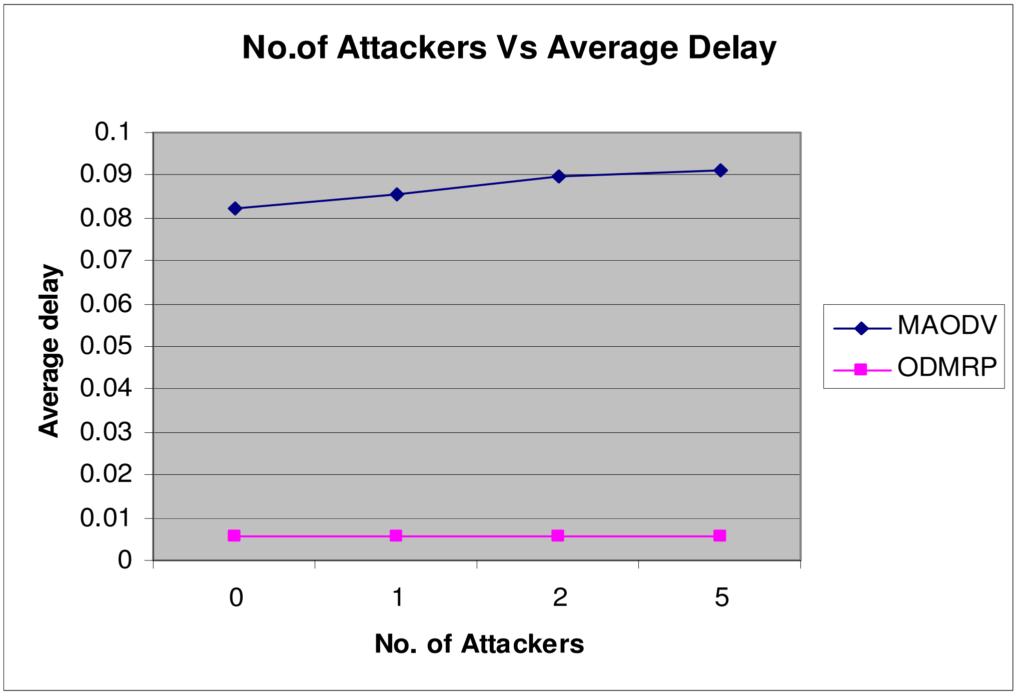


FIGURE 9: Variation of Average Delay with attackers

Figure 9 shows the variation of end to end delay for different numbers. There seems to be an increase in the delay in the presence of attackers. Also the delay is more in MAODV than in ODMRP. This is due to the fact that non shortest paths containing black hole nodes are selected for routing the packets.

5. CONCLUSION

Security is one of the major issues in MANETs. In this paper the effect of black hole attack on MAODV and ODMRP are analysed and compared under different scenarios.The performance of a multicast routing protocol under black hole attack depends on factors such as number of multicast senders, number of multicast receivers and number of black hole nodes

From the simulation results it is observed that, the packet delivery ratio reduces with increased mobility of the nodes and also with increased number of black hole nodes and affect the performance of the network. Also the packet delivery ratio is higher for large number of receivers for the same number of attackers. That is, the effect of the attack is more in a small group than in a large group. A large group is able to withstand the attack to a reasonable extent when compared to a smaller group which is easily susceptible to attacks. This can be attributed to the existence of alternate paths for routing the data packets.

The results also depict that the delay increases with increase in group size and increase in number of attackers. This is because of the fact that non shortest paths containing black hole nodes are selected for routing the packets.

When comparing the performances of MAODV and ODMRP under black hole attack, a general conclusion is that, given the same number of attacker nodes, a mesh based protocol like ODMRP outperforms a tree based protocol like MAODV. This is because to the fact that redundant routes in the mesh of ODMRP provide alternate paths for data delivery.

The simulation results and analysis may pave way for researchers to propose solutions to counter the effect of black hole attacks thereby improving the network performance. Given the constrained resources and the rapidly varying conditions in which the nodes operate, any authentication mechanism that can prevent malicious nodes from participating in the routing process and identify secure routes may provide a proper solution to tackle black hole attack.

6. REFERENCES

1. D. Djenouri, L. Khelladi and N. Badache, A Survey of Security Issues in Mobile Ad Hoc and Sensor Networks, IEEE Communication Surveys & Tutorials, Vol. 7, No. 4, 4th Quarter 2005.
2. L. Zhou and Z. J. Haas, Securing Ad Hoc Networks, IEEE Network Magazine, Vol. 13, No. 6, Nov./Dec. 1999, pp. 24–30.
3. P. Papadimitratos and Z. J. Haas, Secure Routing for Mobile Ad hoc Networks, Proceedings of Communication Networks and Distributed Systems, Modeling and Simulation Conference (CNDS'02), San Antonio, Texas, Jan. 2002, pp. 27–31.
4. E. A. Mary Anita and V. Vasudevan, Performance Evaluation of Mesh based Multicast Reactive Routing Protocol under Black Hole Attack, IJCSIS, Vol. 3, No.1, 2009.
5. S.Lee, M.Gerla and C.Chain, On Demand Multicast Routing protocol-(ODMRP), Proc. of the IEEE Wireless Communication and Networking Conference (WCNC), September 1999
6. A. Vasiliou and A. A. Economides, Evaluation of Multicasting Algorithms in Manets, PWASET, vol. 5, April 2005, pp. 94-97.
7. H. Deng, W. Li, and Dharma P. Agrawal, Routing Security in Ad Hoc Networks, IEEE Communications Magazine, Special Topics on Security in Telecommunication Networks, Vol. 40, No. 10, October 2002, pp. 70-75.
8. Al-Shurman, M. Yoo, S. Park, Black hole attack in Mobile Ad Hoc Networks, ACM Southeast Regional Conference, 2004, pp. 96-97.
9. Sanzgiri K., Dahill B., Levine B. N., Shields, C., and Belding-Royer, E. M., Authenticated routing for ad hoc networks, IEEE Journals on Selected Areas in Communications, 23(3), 2005, 598- 610.
10. B. Sun, Y. Guan, J. Chen and U. Pooch, Detecting black hole attack in mobile ad hoc networks, Personal Mobile Communications Conference. 2003.5th European (Conf. Publ. No. 492), pp. 490 – 495, April 2003.
11. Pankaj Kumar Sehgal & Rajender Nath, A Encryption Based Dynamic and Secure Routing Protocol for Mobile Ad Hoc Network, International Journal of Computer Science and Security (IJCSS), Volume (3) : Issue (1) 16
12. A.Patcha and A.Mishra, Collaborative security architecture for black hole attack prevention in mobile ad hoc networks, Radio and Wireless Conference, 2003. RAWCON '03, Proceedings, pp. 75-78, 10-13 Aug. 2003.
13. E. A. Mary Anita and V. Vasudevan, Black Hole Attack on Multicast Routing Protocols, Journal of Convergence Information Technology, Vol.4, No.2, pp.64–68, 2009.
14. C. Siva Ram Murthy and B. S. Manoj, Ad hoc Wireless Networks- Architectures and Protocols, Pearson Education, 2007

.

COMPUTER SCIENCE JOURNALS SDN BHD
M-3-19, PLAZA DAMAS
SRI HARTAMAS
50480, KUALA LUMPUR
MALAYSIA